

Kapitel 9: Bildungsstandards und externe Überprüfung von Schülerkompetenzen: Mögliche Beiträge externer Messungen zur Erreichung der Qualitätsziele der Schule

Web-Dokument 9.5

Globale Vergleiche zwischen high-stake- oder low-stake-Systemen

Herbert Altrichter und Anna Kanape-Willingshofer

Die Unterscheidung zwischen high-stake- oder low-stake-Systemen hat eine Reihe von Untersuchungen nach sich gezogen, die sich die globale Frage stellen, ob und welches der Systeme zu einer Leistungssteigerung und zu anderen Systemzielen beiträgt. Zunächst bringt dieses Systemmerkmal offenbar unterschiedliche Lernkulturen hervor: Pedulla (2003) hat in einer Befragung von 4 200 US-amerikanischen Lehrer/innen signifikante Unterschiede zwischen high stake-Systemen und low stake-Systemen in verschiedenen Dimensionen (z.B. Schulklima, Abstimmung der Unterrichtsmethoden auf die Testungen, allgemeine Aufmerksamkeit für Bildungsthemen) festgestellt.

Die Frage nach eventuellen Leistungsvorteilen eines Systemtyps geht meist nach einem einfachen bildungsökonomischen Produktionsmodell davon aus (vgl. Bishop, 1995; Bishop & Wößmann, 2004), dass Akteure ihre Anstrengungen nach einem Kosten-Nutzen-Modell kalkulieren: Wenn Prüfungsergebnisse mit gravierenden Konsequenzen verbunden sind, dann zeigen Schüler/innen erhöhte Lernleistungen; wenn Prüfungsergebnisse auch Konsequenzen für Schulen und Lehrpersonen haben, dann werden deren Ressourcen gezielter für (prüfungsrelevantes) Lernen eingesetzt. Die Ergebnisse dieser Globalvergleiche zwischen stark und schwach sanktionierenden Systemen (für die meist Unterschiede zwischen US-Bundesstaaten oder zwischen PISA-Teilnehmerländern genutzt werden) sind keineswegs eindeutig und auch aufgrund von verschiedenen Variablendefinitionen, Datenqualitäten und Zeiteffekten (z.B. Welches Prüfungsregime herrschte vor der Einführung einer stark sanktionsbasierten Reform? Wie lange war die Wirkungszeit?) schwer interpretierbar. Maier (2010, S. 142) hat eine solche Interpretation dennoch unternommen und kommt zu folgendem Ergebnis: „Bei relevanten Testkonsequenzen für Schüler und Schulen steigen die Leistungen, allerdings steigt auch die soziale Selektivität. Häufigeres Testen [auch ohne gravierende Konsequenzen; d.Verf.] konnte dagegen die soziale Selektivität des Lehrerurteils in der Sekundarstufe I reduzieren. [...] Von besonderer Bedeutung sind jüngere Studien auf Basis von Indexvariablen, die [...] relativ valide zeigen, dass Rechenschaftsdruck auf Schüler- und Schulebene zu neutralen bis positiven Effekten auf Schülerleistungen führen kann. Diese positiven Effekte beschränken sich jedoch immer wieder auf einzelne Fächer (Mathematik) oder Schulstufen (z.B.

Sekundarstufe). Die Hinweise auf eine Reduktion oder eine Erhöhung der sozialen Selektivität sind insgesamt uneinheitlich.“

Referenzen:

Bishop, J. H. (1999). Are national exit examinations important for educational efficiency? *Swedish Economic Policy Review*, 6, 349-398.

Bishop, J. H. & Wößmann, L. (2004). Institutional Effects in a Simple Model of Educational Production. *Education Economics*, 12 (1), 17-38.

Maier, U. (2010). Effekte von testbasiertem Rechenschaftsdruck auf Schülerleistungen: Ein Literaturüberblick zu quasi-experimentellen Ländervergleichsstudien. *Journal of Educational Research Online*, 2 (2), 125-152.

Pedulla, J. J. (2003) State-mandated testing: What do teachers think? *Educational Leadership*, 61 (3), 42-46.